AD_____

Award Number: DAMD17-98-1-8255

TITLE: High-Throughput Array Instrument for DNA-Based Breast Cancer Diagnostics

PRINCIPAL INVESTIGATOR: Harold P. Swerdlow, Ph.D.

CONTRACTING ORGANIZATION: Karolinska Institute
S-171 77 Stockholm, Sweden

REPORT DATE: October 2000

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
Distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 2000	Annual (1 Oct 99 - 30 Sep 00)

4. TITLE AND SUBTITLE
High-Throughput Array Instrument for DNA-Based Breast Cancer Diagnostics

5. FUNDING NUMBERS
DAMD17-98-1-8255

6. AUTHOR(S)
Harold P. Swerdlow, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Karolinska Institute
S-171 77 Stockholm, Sweden

E-MAIL:
harold.swerdlow@cgr.ki.se

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; Distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

The aim of this project is to develop improved methods for the production, processing, and analysis of microarrays (gene-chips) for breast-cancer diagnostics. These methods are based upon large numbers of discrete DNA spots placed on glass microscope slides typically, and hybridized to a probe derived from a tissue or blood sample. For breast-cancer diagnostics, there are two likely predominant modes of use. Tissue samples from suspect lesions, e.g., can be assessed for patterns of gene expression on cDNA arrays, indicative of a specific disease type or state. Additionally, analysis of known mutations can be performed on oligonucleotide-based arrays.

We are optimizing cDNA microarray methods, and have achieved excellent limits of detection, although not good enough to detect the rarest transcripts. To improve this situation, we are redesigning our novel imaging (non-scanning) fluorescence microarray reader that currently performs four times better (with respect to limit of detection) than a commercial scanner, and does so about 300 times faster.

Our next goal is to implement oligonucleotide arrays by various methods including both padlock probe-based ligation assays and allele-specific hybridization tests for specific known mutations. We are also actively improving direct covalent attachment of oligonucleotides to glass, and methods for printing oligonucleotide arrays much faster and cheaper.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer, arrays, chips, DNA diagnostics, fluorescence imaging, ligation	12
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

Introduction

This project involves improvements in chip and microarray-based technology for breast-cancer diagnostics. Two different applications can be considered for this technology. Oligonucleotide-based allele-specific mutation detection can be used for identifying known mutations in the breast-cancer susceptibility genes, BRCA1 and BRCA2. We have proposed to use padlock probes (1) for this purpose, taking advantage of the exquisite specificity of DNA ligases to discriminate against mismatched bases as a way to probe specific mutations. Secondly, tissue biopsies and other patient materials can be subjected to the methods of cDNA-based microarrays for analyzing mRNA expression levels in the methods derived from Patrick Brown's laboratory (2). These methods have a great potential to aid in performing pathology on suspected breast-cancer lesions, for example.

Our major focus is on instrumentation to perform microarray analysis easier, cheaper, and with higher sensitivity. The last issue is extremely important, as the use of microarrays for mRNA expression analysis is hindered by the fact that the methods are not sensitive enough to detect low copy number transcripts. For example, in an analysis we did, using previously uncharacterized genes, only about 1/3 - 1/2 of the genes were detectable using microarrays.

Although we are in the second year of this project, we still have not spent more than half a years worth of funding, as we have concentrated on setting up the cDNA methods using institutional funding. However, this coming year we expect to accelerate our pace of research, and move boldly forward.

Body

Our two main areas of focus this past year have been 1) to further develop cDNA array methodology, i.e. optimization of methods to improve reproducibility, uniformity, reliability, and signal-to-noise (aim 2, Statement of Work - SOW), and 2) to refine the imaging (non-scanning) fluorescent microarray reader we have designed (aim 4, SOW).

1) cDNA Arrays

As we reported in our 1999 Annual Report, we have been able to import the technology of Pat Brown's laboratory (http://cmgm.stanford.edu/pbrown/protocols/) and make some improvements and optimizations. The goal was to increase sensitivity to see even weakly expressed mRNAs, while providing more consistency day-to-day in our results.

A key consideration in microarray methods is spot-to-spot reproducibility which impacts both the statistical significance of data, and the limit of detection. We had previously used 0.02% Sarcosyl for achieving greater spot uniformity. Recently, we have followed a suggestion we saw on a microarray e-mail list server, and experimented with using dimethyl sulfoxide (DMSO) as a spotting solution. We compared the following solutions as solvent for the cDNA amplicons to be spotted: 1) 3xSSC with 0.02% Sarcosyl, 2) 50% DMSO, 3) A commercial spotting solution from TeleChem International called Arrayit, and 4) 3xSSC alone. The results are shown in Figure 1. From this figure, it is clear that DMSO works much better than the other solutions. Spotting uniformity is better and overall signal after hybridization is improved.

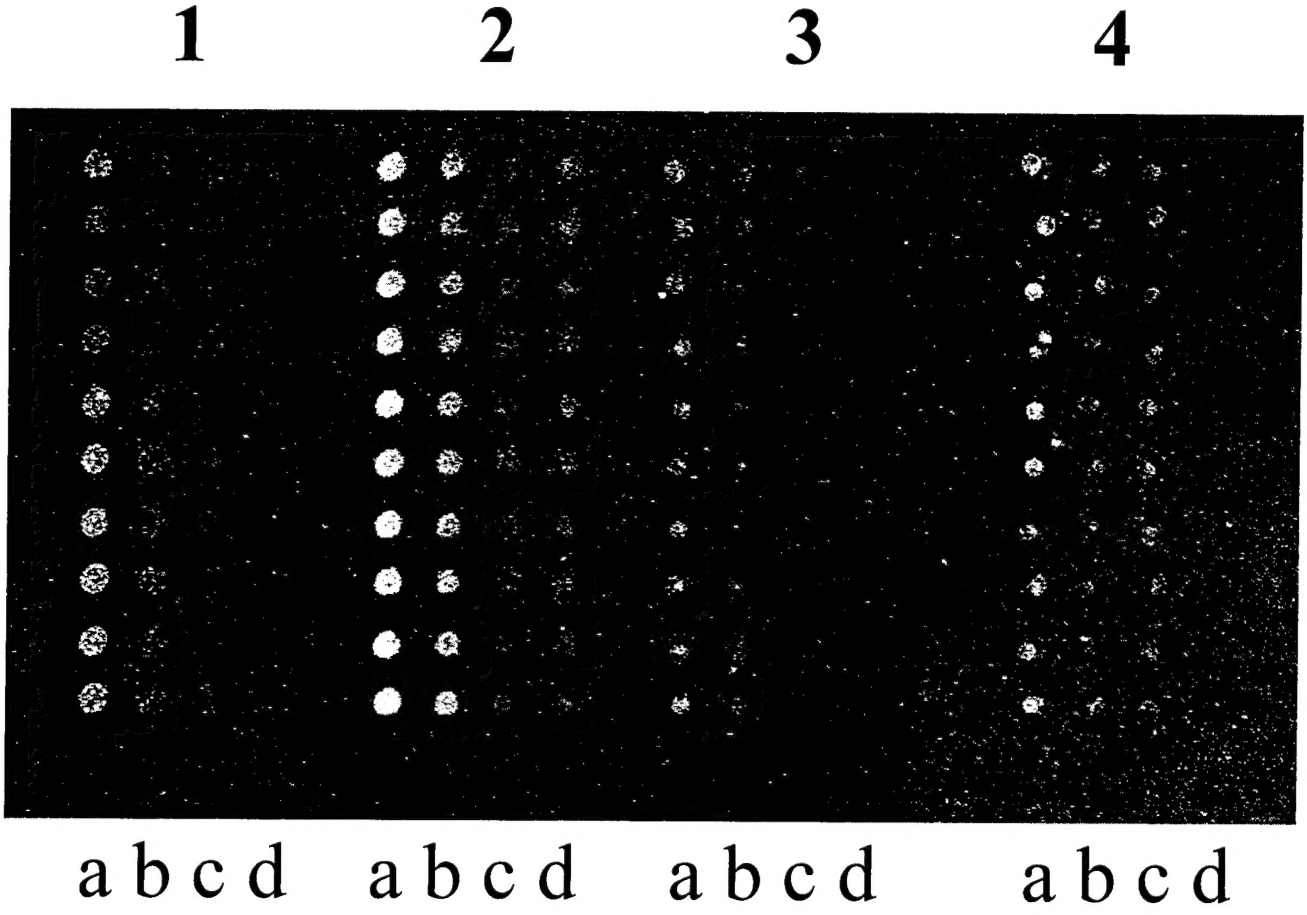

Figure 1. Comparison of spotting buffers. Four sets of spots (human GAPDH) were deposited with different spotting solutions using a Genetic Microsystems' (GMS - now part of Affymetrix) model 417 arrayer, and hybridized to a human brain mRNA labelled with Cy3. After washing, the images were scanned using a GMS 418 reader and the results are displayed. Spots in region 1 were spotted with 3xSSC and 0.02% Sarcosyl; region 2: 50% DMSO; region 3: Arrayit brand spotting solution; region 4: 3xSSC. There are 10 identical spots aligned vertically for each concentration of DNA. Within a set of spots, the DNA concentrations are (L to R), a) 100 ng/ul, b) 25 ng/ul, c) 10 ng/ul, and d) 5 ng/ul.

We also sought to fix the irreproducibility seen with polylysine slides. We have tried many other slides, and still believe that polylysine gives the best signal-to-noise (S/N) ratios of all of the alternatives. However, we have seen a lot of batch-to-batch variation and problems with stability. We have therefore decided to make our own polylysine slides. The procedure is simple (http://cmgm.stanford.edu/pbrown/protocols/1_slides.html), and appears to be much more reliable. The results of a direct comparison are shown in Figure 2. In this figure the home-made polylysine slides gave about the same signal as commercially-available ones. However, when the slides were washed with an improved protocol, we were able to obtain almost twice the signal as either the commercial slides or the ones treated with the standard Brown Lab protocol.

2) Imaging Reader

Our interest in designing a new fluorescent microarray reader derives from the issue of sensitivity. Increased sensitivity is key to making microarray methods more useful. The paradigm we are promoting is that scanning systems are inferior to imaging systems. This has two components. The first is the issue of optical saturation. Scanners can give quite adequate signals when the scan speeds are slow, or the number of pixels is small. As the number of pixels increases, if the total scan time is kept constant, the time spent on each pixel must be reduced. The only way to achieve the same amount of signal is to increase the laser power, which, apart from the expense, has a limitation. At some intensity, all the fluorophores that can be excited are already excited, this is known as optical saturation or ground-state depletion. At that point, increasing the laser power only serves to increase the background noise, and S/N ratios begin to degrade. This translates into a loss of sensitivity for low copy number mRNAs. A second problem with scanning detectors is the dynamic range. Scanners usually have fixed speeds, which means fixed sensitivity ranges in terms of fluors/μm^2. Imaging readers, on the other hand, can take a picture of the entire array or subarray at once. Exposure times can easily be adjusted over a huge range, e.g., from a few milliseconds to many minutes. Therefore the dynamic range can be extended by many orders of magnitude, an advantage for a microarray reader. The most abundant transcripts as well as the least abundant mRNAs should be able to be visualized in the same experiment.

We embody this concept in an evanescent-wave / CCD-camera based system. In our first prototype, an argon-ion laser operating at 488 nm is used to excite Cy3 dyes on a standard microscope slide. The light from the argon-ion laser is focused by a microscope objective onto a standard fiber-optic waveguide. The end of the fiber is used to introduce light into the side of the slide at an angle of 30 - 45°. This light is reflected many times down the length of the slide by total internal reflection. The slide in effect acts as a planar fiber-optic waveguide. The light also spreads in the lateral dimension providing a reasonably uniform beam through the slide. A quantum-mechanical phenomenon known as evanescent-wave excitation works to produce fluorescence on the surface of the glass slide. Briefly, light contained in the core of the slide is totally internally reflected, i.e., trapped completely inside. However, the energy field penetrates a few nanometers into the surrounding medium. This energy can be used to excite molecules located close to the surface, and has been used for various high-sensitivity detectors (3). The light can be collected normal to the slide's surface, in which direction no background excitation energy can propagate.

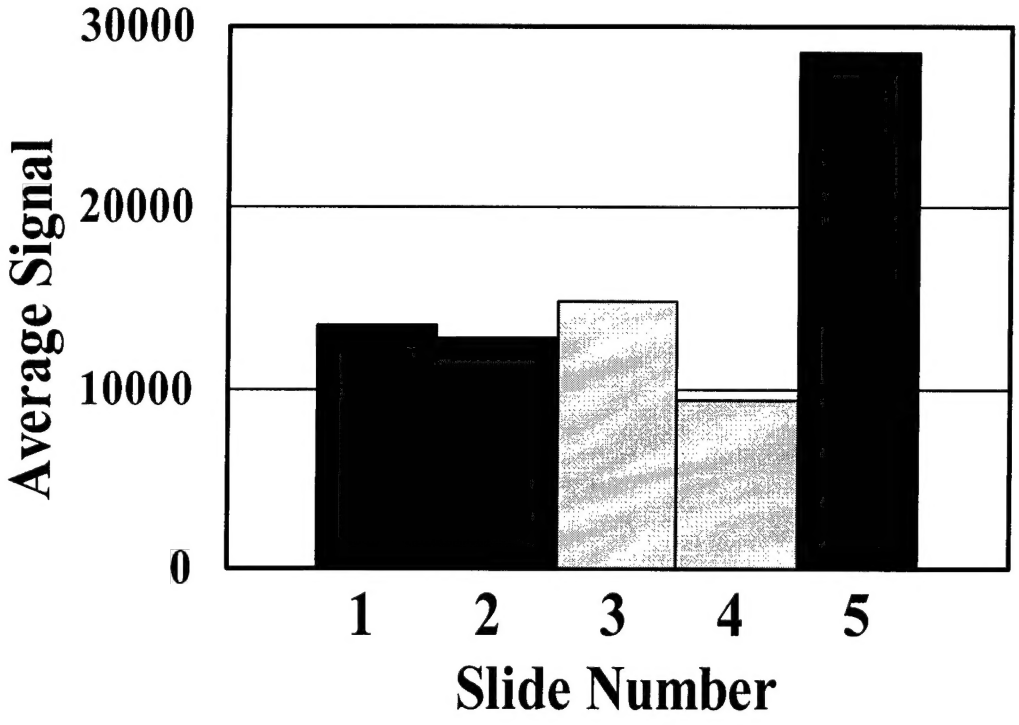

Figure 2. Comparison of various slide surfaces. Average signal for 15 spots is shown for the following slides: 1&2: commercial polylysine (Menzel-Glaser, Germany), 3&4: home-made polylysine, and 5: polylysine coated after improved wash procedure. Spots were GAPDH-derived PCR products. Hybridization was to a human brain mRNA sample.

The light emerging from the slide is collected by a high numerical aperture, wide-field optical train, consisting of two large diameter lenses surrounding a very efficient holographic notch filter. This lens system collects a large amount of light, and from a wide area (1-2 cm) on the array. Emitted light is imaged onto a CCD camera, which collects half a million pixels per frame, thereby giving both wide-field imaging and high spatial resolution.

In our last report, we were able to show that this prototype could read microarrays in about 1 sec. at a limit of detection of 1 nm dye, corresponding to 30,000 fluors in a spot, 800 fluors in a pixel, or 2 fluors/μm^2. Although I believe these numbers are impressive, especially the speed, which is more than two orders of magnitude better than our commercial scanner, the sensitivity is still not much of an improvement over conventional scanners. Therefore, we spent the last year refining this detection strategy, hoping to lower the sensitivity limits.

- Sample / dye issues

We investigated the behavior of Cy3 and other dyes with respect to drying. We found that Cy3 performed far better than fluorescein when dried. However, there still was a significant loss of signal for Cy3 upon drying. Our recent results indicate that there is a maximum signal during this process, presumably due to an increase in concentration before complete drying occurs. To take advantage of this observation would be difficult in practice, as drying kinetics for 50-100 pl drops are not something easy to control. An alternative strategy would be to perform scanning in the wet state, and although this is certainly possible in the abstract, we have not yet been tempted to implement such a modification.

- Instrumentation issues

The angle between the incoming beam and the slide was varied to try and optimize the S/N ratio. This measurement is a bit tricky, as the coupling has to be re-optimized for each angle. Nonetheless, we determined our optimum coupling angle to be 35 degrees. We also tried to use a 45° quartz prism to couple light into the slide from the slide's flat surface - a fairly well-established coupling strategy. However, we saw no increase in signal with respect to background.

We experimented with index-matching fluids to improve the coupling efficiency between the fiber-optic excitation waveguide and the slide or coverslip (see below). With an n= 1.44 matching fluid, coupling was visibly improved. However, this oil was a bit too thin to allow the system to be very stable, and we tried other fluids, including a polymer solution, with not much success.

The excitation fiber had some visible background fluorescence. We stripped the plastic cladding from the fiber at the ends to remedy this. We also experimented with a variation of the system using a mirror and a focusing microscope objective to introduce the light directly from the laser beam without a fiber, with no significant improvement. We replaced this microscope objective with a synthetic fused-silica (quartz) singlet lens in an effort to reduce background. We also removed all the optics, and pointed the laser directly onto the edge of the slide. Lastly, we replaced the evanescent-wave excitation mode with a direct illumination strategy. Although we saw no significant improvement, the direct

illumination strategy appears to be as good as the evanescent wave excitation scheme, which of course is more complicated.

Another variation we experimented with, was to use the quartz prism itself as a substrate for spotting the arrays. In this mode, light is bounced just once off a face of the prism by total internal reflection. This same face is spotted with dye, and the evanescent wave is used to excite the dye on that face. Coupling efficiency is far better in this approach; however, the light only bounces once, and it appears that the performance of this technique is no better than the other methods.

We investigated whether changing the laser-excitation wavelength would provide an improved S/N for the Cy3 dye. We obtained filters compatible with many laser wavelengths and tried both a green helium-neon laser as well as a frequency-doubled neodymium-YAG laser. This gave us access to the following wavelengths: 488, 514, 528, 532, and 543 nm. Although the higher wavelengths gave somewhat better signals for the same laser power, the improvement was not terribly significant.

We investigated whether or not the collection lenses were causing problems for our system. They did not appear to be a significant source of background fluorescence.

An additional holographic notch filter was added to the system. However, we saw no improvement. This implies that most of the background we see is due to fluorescence (or possibly Ramen scatter), and not to laser (Rayleigh) scatter.

- Slide issues

Standard glass slides are certainly not optimal for evanescent-wave-based detection strategies. Some groups have gone as thin as 1 μm planar waveguides to increase the intensity of illumination (at a fixed excitation power), while background remains roughly constant. We first tried ordinary glass coverslips for preparing arrays. S/N is improved by a factor of about 10, corresponding to the reduction in thickness from 1 mm to 100 μm. However it becomes much harder to perform proper side-on coupling.

We tried a few different versions of quartz slides as well. The first quartz slides we tried were standard natural quartz, which were not at all satisfactory. Subsequently we switched to synthetic fused-silica, a man-made, ultra-pure grade of quartz. The background was much lower; limits of detection improved minimally.

Slide cleanliness was found to play a major role in background. For several different solvents, we investigated the purity (assayed by fluorescence spectrometry) and residue left after drying. We found a workable, if not globally optimal procedure. We wash with acetone and ethanol to remove organic films. However, since both of these solvents leave residues, we follow this treatment with a soap wash, and a rinse in high-quality distilled water. This procedure reduces the background fluorescence to tolerable levels. We also experimented with other strategies, including acid and base washes, and other combinations.

- Summary

All the above changes have taught us a lot about the non-imaging microarray reader system. The bottom line is somewhat better than before. Background fluorescence in the slides still gives a large background. This seems generally true for both the evanescent-wave excitation and the direct illumination schemes, since the entire volume of the slide is illuminated with excitation light. Presumably this points out one inherent advantage of confocal schemes (which cannot be used in imaging mode) for this application. Nonetheless, our concentration detection limit is currently about 500 pM, or 1 fluor/μm^2 - about 4x better than our commercial scanner in our direct experiments (the manufacturer's claims notwithstanding). However, our system is still much faster than commercial scanners. This may have a great advantage for real-time microarray applications in the future.

Key Research Accomplishments

- optimization of spotting buffer for cDNA microarrays
- incorporating the home-made polylysine procedure
- improving the non-scanning (imaging) detection strategy for microarrays
- achieving detection limits of 0.5 nM dye in a time of about 1 sec.

Reportable Outcomes

- none this period

Conclusions

The advent of microarrays has produced a wealth of new information, some good, some bad. With the ability to perform thousands of hybridization-based assays simultaneously on the microchip format, comes the realization that bad science can also be done thousands of experiments at a time. For example, the recent results from Lander's group, discriminating acute myeloid leukemia (AML) from acute lymphoblastic leukemia (ALL)were quite impressive (4). However, the temptation to apply these concepts to solid tumors must be evaluated with care. The largest problem is the presence of other contaminating tissues in the samples, for example vascular tissue, and "contaminating" normal cells in tumor samples. These latter problems make breast cancer diagnosis and prognosis by microarray methods a tricky proposition. That said, the potential for accurately identifying breast tumor biopsies and especially for discovering genetic sub-class distinctions heretofore unknown by phenotype/morphology can not be ignored. Thus we feel confident that the future of applying microarrays to the field of cancer diagnostics and prognostics is a promising one indeed.

A key issue is the difference between oligonucleotide-based methods and those based upon cDNA clones/PCR products arrayed on a chip. The latter methods are becoming very popular due to the fact that oligonucleotides are difficult to prepare in-situ, and may have

problems with the yield-per-step, and certainly have problems with sequence-specific differences in both accessibility for hybridization and in melting temperature (Affymetrix technology). Another problem is the lack of specificity when performing genome-wide scans for short oligonucleotides. Lastly the chips available from Affymetrix are currently quite expensive and therefore prohibitive for large studies. These problems and the fact that we have spent a large amount of time optimizing cDNA methods in our laboratory notwithstanding, we are still quite committed to oligo-based approaches for microarray diagnostics.

We favor two methods for using oligonucleotides on microarrays. Allele-specific experiments with longer oligonucleotides are currently being planned here. Under the direction of DR. Zicai Liang (co-director of the array core facility), we will test 30-70 mers for specificity, temperature independence of hybridization, and compare these results with the relative costs. Secondly, as stated in the original application, we will look into padlock probes - a method invented by Ulf Landegren and his colleagues (1) - as a way to use the specificity of ligases to measure mutations in breast-cancer specific genes. This latter project will be more fruitful, if our department hires a senior postdoctoral fellow from Dr. Landegren's lab; this is currently under serious negotiation.

References

1) Nilsson, M., Malmgren, H., Samiotaki, M., Kwiatkowski, M., Chowdhary, B.P., and Landegren, U. (1994) Padlock probes: circularizing oligonucleotides for localized DNA detection. Science 265, 2085-2088.

2) Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P.O., and Davis, R.W. (1996) Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. Proc. Natl. Acad. Sci. USA 93, 10614-10619.

3) Plowman, T.E., Reichert, W.M., Peters, C.R., Wang, H.K., Christensen, D.A., and Herron, J.N. (1996) Femtomolar sensitivity using a channel-etched thin film waveguide fluoroimmunosensor. Biosensors and Bioelectronics 11, 149-160.

4) Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D., and Lander, E.S., (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286, 531-7.